Robust photometric invariance in machine color vision
Stokman, H.M.G.

Citation for published version (APA):
Stokman, H. M. G. (2000). Robust photometric invariance in machine color vision

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 1

Introduction

1.1 Machine Color Vision

When asking a passer-by in Amsterdam for directions, the answer is either "I don't know" or else you are told the precise route. When asking the same question in Jakarta, Indonesia, the people in the street are eager to help and send you off in some direction, regardless of whether they actually know your destination. In contrast to the exact type of answer given in Amsterdam, directions given in Jakarta come with an amount of uncertainty. It is therefore good practice in Jakarta to pose the same question several times, thereby obtaining an estimate of the certainty what direction to follow. However, rather than asking the question several times, suppose it is possible to obtain an expression for the probability that the given direction is false, would this help in finding your destination?

When a machine performing a visual inspection task is asked for the color of a red mobile phone, say, the answer also comes with some limited certainty as sensor measurements are typically subject to uncertainty. Suppose that it is possible to obtain an expression for the probability that the color measured by the camera is wrong, would this help establishing the color of the phone? Perhaps, due to the orientation of the phone with respect to the camera, one side reflects more light than the other side. As a result, at one side a more brightly color red is measured than at the other side. Alternatively, perhaps an error occurred during the dying of the phones. Should the production of red phones be halted immediately to prevent further waste? Suppose that it is possible to decompose the observed color into a color component (red) and a component encoding the amount of reflected light, does this help answering the question of whether the color of the phone corresponds to the target color?

From the example it can be abstracted that color is important in industrial inspection. Automated visual inspection, further elaborated upon in [9] [8] [38] [53] [12], has the potential of providing industry with two advantages. The first advantage is a possible increase of the quality of products, the second advantage is a possible decrease in the price of the product. Readers and advertisers of magazines and news-
papers constantly demand more color [50] [49] and the market does not only demand more quantity, but also a better quality of color [34]. Color is an especially important factor in apparel and food inspection where changes in shade or variations in color indicate the quality of the food [11]. An advantage of electronic evaluation of color is that humans are not reliable color inspectors, primarily because humans do not have a very good memory for color [38]. Therefore, machine color vision appears as a legitimate technique for visual inspection applications in, e.g., agricultural, textile, and packaging industry.

In this thesis, digital color images are used for visual inspection tasks. As the answer is subject to uncertainty, in the thesis we aim to estimate the quantity of uncertainty of the answer (When asking for directions, what is the probability that the answer is wrong?) We concentrate on sensor noise and investigate whether such estimate improves the performance of the inspection task (When asking for directions, does the extra information on the probability that the answer is wrong actually help to find the destination?) Further, the answer may also be subject to photometric effects of shading, highlights, geometry of the object, viewing direction, and to the color and intensity of the illumination (The color of the red phone appears to have changed, is it because of errors in the dying process, or is it due to a change in the orientation of the phone?) This can be seen as a source of error, an unwanted effect, or a wanted effect. At any rate, in addition to estimating sensory uncertainty, in the thesis an attempt is made to derive expressions for the invariance with respect to one or more of the above mentioned photometric effects.

In the next section, the use of color in machine vision is discussed. In section 1.3, various color models are examined briefly. In section 1.4, the issues just raised are described in more detail An overview of the thesis is given in the context of these issues in section 1.5.

### 1.2 Taxonomy of Machine Color Vision Systems

In this section, a taxonomy is given of various color imaging systems. Monochrome vision using colored illumination, red-green-blue (RGB), and multispectral cameras are discussed.

#### 1.2.1 Machine Color Vision by Monochrome Cameras

Factors such as light source intensity and angle of illumination are crucial for accurate color monitoring [11] [10]. Moreover, an observed color does not only depend on the color of the object but also on the the sensitivity of the color filters of the sensor and on the color of the light source [63]. An often encountered task in computer vision is the separation of a foreground object from the background. Carefully choosing colored illumination or a color filter [30] [31] with respect to the objects in the scene can simplify the inspection tasks. When the foreground and background have different colors, the use of color filters may increase the contrast between the two. Such method can be applied for images taken by monochrome as well as red-green-blue (RGB)
The quality of lighting appears to have received much more attention from industrial developers [57] [59] [32] [42] [1] than from the vision research community. For practical applications, it is safe to conclude that it is almost always cheaper to improve the lighting than the image processing [4]. The illumination may imply the difference between success and failure of the development of an automated visual inspection system. For example, the use of backlighting may provide a high contrast between an object and the background, the use of diffuse front lighting may eliminate shadows, and polarized lighting may remove specular reflections [13] [40]. However, it may not always be possible to tailor the light by any of these methods, for example when the scene is outdoors. In these cases, it is advantageous to obtain image processing results invariant to photometric effects of the angle of illumination, light source intensity, shading, surface geometry, etc. [15] [3].

1.2.2 Machine Color Vision by Red-Green-Blue Cameras

An advantage of using RGB color compared to monochrome images is that RGB images convey three times more information than monochrome images [43] [47] [11]. More information has the potential to increase the precision of machine vision algorithms [37]. RGB images also allow to distinguish objects having equal image intensities based on their difference in color.

Another advantage of color images was recognized in the mid-1980s by the image processing community when color variations of surfaces were modeled as a physical process. A landmark was reached in the series of articles collected by Wolff, Shafer and Healey [62]. The physics-based color image processing was initiated in 1985 by Shafer [46] who presented the dichromatic reflection model. Klinker [29] developed a color segmentation algorithm based on the reflection model. The approach is useful because algorithms can account for the physical model of reflection. Klinker's algorithm segments the image into connected regions corresponding to surfaces of a single material. It is more sophisticated than other segmentation techniques [44] [41] [7] [36], in that it is designed not to break the segmentation results because of highlights or because of intensity changes due to geometrical variation in the scene. Using the dichromatic reflection model, Tsang and Tsang [56] show that edge detection in the hue color space is effective in suppressing specular reflection. Lee [33] uses the reflection model to detect the color of the illuminant. It is therefore of interest to investigate how the dichromatic reflection model can be used to derive (in theory) methods for the visual inspection of objects, invariant to factors such as illumination, scene geometry, surface reflectance and material properties.

1.2.3 Machine Color Vision by Multispectral Cameras

An advantage of using multispectral images compared to RGB images is that multispectral images convey more information than RGB images [5] [24]. More information has the potential to increase the precision of machine vision algorithms [21]. Multispectral images also allow to distinguish objects having equal RGB intensities based
on their spectral differences. For example, the problem of metamerism, where the same three dimensional color coordinate corresponds to several different spectra can be avoided using multispectral data [19]. For applications where, e.g., inks are mixed to obtain a target color, multispectral data may also provide the possibility to trace the cause of an occurring color error.

Multispectral imaging has received a great deal of attention as such in recent years. Spectral measurements are used, e.g., in remote sensing [16], computer vision, and industrial applications [25]. Spectral images are obtained, for example, by a CCD-camera with narrow-band interference filters [27]. Tominaga [54] [55] describes two generations of a multichannel vision system based on the use of a CCD-camera and six color filters to reconstruct the surface spectral reflectance and illuminant spectral power distribution. Baronti et al. [2] used a multispectral imaging system for the non-invasive analysis of works of art. Haneishi et al. [18] designed five color filters for archiving spectral images of art works.

Multispectral data complicates image processing compared to RGB images. First, it increases the amount of required processing. Moreover, there is the problem of acquiring multispectral data. Several approaches exist. A two-dimensional image can be obtained at different wavelengths at different times. Alternatively, a one-dimensional line image acquires the reflectance at different wavelengths simultaneously. The former approach is implemented by employing, e.g., a rotating filter wheel in front of a monochrome camera. This two-dimensional imaging at one wavelength at a time suits to laboratory applications with a stationary sample. A drawback of the approach is the potential occurrence of chromatic abberation. The latter approach is implemented by an imaging spectrograph coupled to a monochrome camera. One dimension of the camera (spatial axis) records the line pixels and the other dimension (spectral axis) the spectral information for each line pixel. The approach is appropriate for a moving target or a moving sensor. Instruments also exist which are mounted to a monochrome CCD-camera and which contain N different color filters. The N differently filtered 2-D images are then projected onto the monochrome CCD-grid. Increasing the spectral resolution N thus decreases the spatial resolution. If combined with, e.g., the Kodak blue enhanced KAF-1401E chip to overcome the insensitivity of CCD-cameras in the 400-450 nm wavelength region, these multi-spectral sensors enable spatially resolved multi-spectral imaging. A taxonomy of these systems is given in table (1.1). Thinking beyond the current limitations of the technology, it would be desirable to acquire two-dimensional images at a single moment of both high spatial resolution and high spectral resolution. Such systems are not available yet.

Summarizing, color is important for inspection in the agricultural, textile, packaging and advertising industry. Therefore, it is of commercial interest to investigate the use of machine color vision. Apart from using traditional RGB cameras, color can also be exploited for monochrome vision. Using multispectral data is of interest because multispectral cameras are relative newcomers in industrial inspection and because multispectral data has the potential of high measurement accuracy. The dichromatic reflection model (in theory) may allow to derive methods invariant to photometric effects. It is therefore of interest to extend the model to data obtained by multispectral image sensors.
1.3 Color Models

In the previous section, different methods were described to obtain sensor data for color vision. For various reasons, the sensor data is often converted into other color spaces. Among one of the oldest color spaces is the 1931 RGB color space \([63]\). The \(XYZ\) tristimulus values are obtained using the color-matching functions of the human observer. The CIE \(L^*a^*b^*\) space is obtained by a non-linear transform of the \(XYZ\) color space \([63]\). The CIE \(L^*a^*b^*\) color space is very important due to its widespread use in industry. The color space is (approximately) perceptually uniform. This means that two colors that are just noticeably different for a human observer correspond to a single numerical difference in color values, regardless of what two colors are compared.

Colors differ in saturation, hue and intensity. Many transforms exist, e.g., \([35]\) \([28]\) \([43]\) \([15]\) which transform RGB colors into the hue-saturation-intensity representation. The differences in the equations are caused by trade-offs between the computational simplicity and the perceptual uniform distribution of the colors over the color space axes. The advantage of the representation is its perceptual comprehensiveness.

A two-dimensional color space is obtained by removing the intensity component of colors. The two-dimensional color coordinates are obtained by intersecting a three-dimensional line through the origin and a RGB color value with a surface. The shape and orientation of the surface can be arbitrary, the only requirement is that a single RGB color value appears as a single intersection with the surface. This way, a chromaticity diagram is obtained. Perhaps the most widely used chromaticity diagram is the \(xy\) color space obtained from the \(XYZ\) space. Via a perspective transformation of the diagram, the \(uv\) diagram is obtained which is approximately uniform. Another well-known chromaticity color space is the normalized red-green-blue color space derived from RGB sensor data. The advantage of the chromaticity diagram representation is its insensitivity to intensity changes.

Yet another color space often encountered is the opponent color space, originally
proposed by Hering [23]. Using the model, it is possible to explain some perceptual color phenomena inexplicable by classical trichromatic theory. Television color spaces are in fact opponent color spaces, because they define an intensity component and two chrominance components. The television color spaces are designed to minimize the bandwidth required to broadcast the color signal. Because the human visual system is less sensitive to details in chrominance than in luminance, the bandwidth of the chrominance signal can be significantly smaller than for the intensity signal. In this thesis, color spaces are investigated for photometric invariance and for sensitivity to sensor noise.

1.4 Problem Statement

The methods described in this thesis are derived from the physics-based dichromatic reflection model. This makes it possible to characterize what kinds of images are likely to be segmented successfully by our algorithms which are based on the model. The model describes the reflection of materials which are optically inhomogeneous. In that case, the light interacts with a medium that causes the bigger component of surface matter, as well as with particles of a colorant to produce scattering and coloration. Many common materials have these characteristics, including most paints, varnishes, paper, ceramics, and plastics, etc. Therefore, we anticipate that images of objects of these materials are properly processed by the proposed methods. As a consequence, the proposed methods are unsuited for homogeneous materials.

The dichromatic reflection model assumes that several conditions hold. Firstly, that there is no inter-reflection among surfaces. Secondly, that there is a single light source, i.e. no diffuse (“ambient”) illumination. Thirdly, it is assumed that the objects are uniformly colored, i.e. the objects have a uniform distribution of the colorants. Furthermore, from the model follows that the color value measured by the camera depends only on the intensity and spectral power distribution of the illuminant, the color of the object, and the sensitivity of the color filters, the orientation of the camera, light source and surface normal, and on whether the object is matte or shiny. In this thesis, we further assume neutral interface reflection, that is that highlights are expected to have the same color of the illuminant. Also, we focus on industrial inspection, where the imaging conditions are typically controlled to some extend. The image processing methods described in the thesis are therefore tested only on images acquired under controlled laboratory conditions.

A uniformly colored object satisfying the dichromatic reflection model may give rise to a broad spectrum of color values due to the dependence of the measured color on the geometry of the object and due to specularities. For many industrial machine vision tasks, it is required to visually inspect objects independent of these photometric effects. To that end, the color data should be transformed into color invariant spaces [17]. However, it was shown by Kender [28] that, given certain input RGB-values, the output values for the standard invariant color models are unstable under the presence of sensor noise. For instance, the essential singularity of normalized coordinates is at black \((R = G = B = 0)\), whereas for the hue the singularity is at
the achromatic axis \((R = G = B)\). As a consequence, both color spaces become unstable near the singularity where a small perturbation of \(RGB\) value might cause a large jump in the transformed values. Traditionally, these effects are either ignored or suppressed by an ad hoc thresholding of the transformed values. For example, Ohta [39] rejects normalized color values if the sum of \(RGB\) is less than 30, and rejects hue values if the saturation times the intensity is less than 9. Healey [22] rejects all sensor measurements that fall within the sphere of radius \(4\sigma\) centered at the origin in \(RGB\) space.

A more elegant strategy to deal with unstable color values would be to generate the reliability of a transformed color together with the output and to incorporate the reliability in the image processing method. An early effort in this direction is the work of Burns and Berns [6] who analyze the error propagation from a measured color to obtain an indication of how the color space transform influences the color values under the influence of noise. Shafarenko, Petrou and Kittler [45] use an adaptive filter for noise reduction in the CIE \(L^*u^*v^*\) space prior to 3-D color histogram construction where the filter size depends on the covariance matrix of the noise distribution in the transformed color space. However, the authors do not compare the results of their adaptive filter with a conventional filter.

The aim of the thesis is threefold:

1. To derive an expression of uncertainty associated with the output of an image processing method.
2. To exploit the obtained expression to improve automatic visual inspection methods.
3. To visually inspect objects using multispectral images made invariant to some or more of the photometric effects described above.

To illustrate this, consider figure (1.4). To visually inspect the objects shown in (1.4a), the object contours may be of interest. The uncertainty associated with the gradient in normalized color space is shown in (1.4b). The uncertainty is employed to distinguish edges caused by sensor noise from real edges. However, edges are also detected due to the occurrence of highlights, object geometry and shading. Employing the dichromatic reflection model, it can be derived that different color spaces are invariant to different photometric effects. The knowledge can be used to classify the object contours as shown in (1.4c). The images are taken from chapter 4 of the thesis.

In the next section, a complete overview of the chapters making up the thesis is given.

### 1.5 Overview of the Thesis

In chapter 2 the obtainable accuracy of color measurement using a spectrograph is investigated. It is investigated if the effect of sensor noise on the computed color difference between a measured color and the target color can be estimated. The advantage for industrial inspection is that the predicted uncertainty is available for
the decision of what action needs to be taken if the measured color difference exceeds a predefined threshold.

In chapter 3, it is investigated if the effect of sensor noise on color invariant spaces can be estimated. The result aimed for is the transformed color value of a pixel with an expression of the reliability of the value. It is then investigated if the information can be used to further improve density estimation \[14\] \[61\] \[48\].

Color edge information from an image can be used to measure or recognize objects in images. False edges are often detected due to sensor noise. These false edges are conventionally eliminated by using a threshold which determines the minimum acceptable gradient modulus. The problem is how to find such threshold value. In chapter 4, it is investigated if the effect of sensor noise on the computation of the gradient magnitude can be estimated, and whether this additional knowledge can be exploited for parameter-free thresholding of color edges. Apart from the occurrence of color edges due to material changes, color edges may also be caused by a variety of photometric effects. The problem is how to distinguish between these different color edge types. Further, it is investigated if detected edges can be classified into shadow, geometry, highlight, or material transition edge.

Chapter 5 contains a study on segmentation based on $K$-means clustering\[20\]. Image segmentation techniques can be used for, e.g., presence verification or for shape or area measurement of objects. The aim is to detect homogeneously colored regions invariant to photometric effects in multispectral images. To this end, different polar angle representations of a spectrum are examined for the dichromatic reflection model. However, such representation may become unstable in the presence of sensor noise, a problem shared by other photometric invariant color spaces. Based on the CCD-
camera sensitivity, a theoretical expression of the certainty associated with the polar angular value is obtained. It is investigated whether the additional information can be used to obtain robust segmentation results under the influence of noise.

For an intelligent system performing industrial inspection while sometimes operating in daylight and sometimes in artificial light, it is required to separate the influence of the changing spectral distribution of the illumination and the object reflectance properties. Therefore, in chapter 6 we obtain estimates of the chromaticity of the light source for conventional red-green-blue images as well as for multispectral images. We apply the clustering technique described in the previous chapter to divide automatically the multispectral image into a number of regions corresponding to uniformly colored objects. Two methods are investigated to obtain the spectral distribution of the illuminant. The result has the potential allow object inspection invariant to the color of the illumination.

Bibliography


